

# An under-approximation for the robust uncertain two-level cooperative set covering problem

Shuxin Ding, Qi Zhang, and Zhiming Yuan

**Abstract**—This paper investigates the robust uncertain two-level cooperative set covering problem (RUTLCSCP). Given two types of facilities, which are called y-facility and z-facility. The problem is to decide which facilities of both types to be selected, in order to cover the demand nodes cooperatively with minimal cost. It combines the concepts of robust, probabilistic, and cooperative covering by introducing “ $\Gamma$ -robust two-level-cooperative  $\alpha$ -cover” constraints. Additionally, the constraint relaxed version of the RUTLCSCP, which is also a linear approximation robust counterpart version of RUTLCSCP (RUTLCSCP-LA-RC), is developed by linear approximation of the constraints, and can be stated as a compact mixed-integer linear programming problem. We show that the solution for RUTLCSCP-LA-RC,  $\varepsilon$ -under-approximate solution, can also be the solution for RUTLCSCP on some conditions. Computational experiments show that the solutions in 333 instances (10125 instances in total) with 12 types which tinely violate the constraints of RUTLCSCP, can be an efficient under-approximate solutions, while the feasible solutions in other instances are proven to be optimal.

## I. INTRODUCTION

The set covering problem (SCP) is one of the most studied combinatorial optimization problems. In the SCP, a set  $\mathcal{I} = \{1, \dots, m\}$  of  $m$  demand nodes, a set  $\mathcal{J} = \{1, \dots, n\}$  of  $n$  potential facility location sites and their building costs  $c_j$  are given. The 0-1 matrix  $A = [a_{ij}]_{m \times n}$  indicates whether a location  $j \in \mathcal{J}$  is able to cover a demand node  $i \in \mathcal{I}$ . The goal of SCP is to find a minimum cost cover of the demand nodes by  $x$  where  $x_j$  is a binary value whether site  $j$  is selected. It is proven to be NP-complete [1].

$$\min \sum_{j \in \mathcal{J}} c_j x_j \quad (1)$$

$$\text{s.t.} \sum_{j \in \mathcal{J}} a_{ij} x_j \geq 1 \quad \forall i \in \mathcal{I} \quad (2)$$

$$x_j \in \{0, 1\} \quad \forall j \in \mathcal{J} \quad (3)$$

The SCP has widely been used in many real-world applications, especially in facility locations [2], where both exact and heuristic algorithms are proposed to deal with it. Daskin [3] considers the facility may not by working with probability, and it can be applied in many applications, e.g., node deployment in wireless sensor networks [4], weapon platforms [5], etc. Beraldi *et al.* [6] proposed the probabilistic set-covering aiming at covering constraint satisfied with a

predefined probability. Aardal *et al.* [7] considered more than one facility type, and proposed a two-level uncapacitated facility location problem. Berman *et al.* [8] first proposed the cooperative cover model with one facility type.

Pereira *et al.* [9] proposed the robust SCP with uncertain cost coefficients within predefined interval. To the best of our knowledge, the robust set covering problem with probabilistic and cooperative covering by two types of facilities has not previously been analyzed. The collaboration of different platforms has been analyzed recently. Xin *et al.* [10] and Xu *et al.* [11] discussed the sensor-weapon-target assignment problem as a collaborative task assignment of sensor and weapon platforms, and Xu adopted Soyster’s robust model [12]. The probability of capturing the target is similar to the cooperative covering in this paper, while the former is regarded as the objective function. We summarize our contributions as follows:

1. A compact mixed-integer linear programming formulation is proposed by utilizing robust optimization and constraint relaxation.
2. The proposed formulation is analyzed on a large set of test cases with 10125 different instances.
3. A majority of the under-approximate solutions are proven to be optimal, while few of them slightly violate the constraints and provide an efficient lower bound.

The rest of this paper is organized as follows. Section II formulates the robust uncertain two-level cooperative set covering problem. Section III presents some properties of the model. Performance evaluation results are presented and analyzed in Section IV. Conclusions are given in Section V.

## II. FORMULATING THE ROBUST UNCERTAIN TWO-LEVEL COOPERATIVE SET COVERING PROBLEM

### A. The Deterministic and Uncertain Two-Level Cooperative Set Covering Problem

In the Two-Level Cooperative Set Covering Problem (TLCSCP), a set  $\mathcal{I} = \{1, \dots, m\}$  of  $m$  demand nodes, a set  $\mathcal{J} = \{1, \dots, n_1\}$  of  $n_1$  potential y-facility location sites and a set  $\mathcal{K} = \{1, \dots, n_2\}$  of  $n_2$  potential z-facility location sites are given. The 0-1 matrix  $A = [a_{ij}]_{m \times n_1}$  or  $B = [b_{ik}]_{m \times n_2}$  indicates whether a location  $j \in \mathcal{J}$  or  $k \in \mathcal{K}$  is able to cover a demand node  $i \in \mathcal{I}$ .  $c_j^1$  represents the costs of building y-facility located at site  $j$ , and  $c_k^2$  represents the costs of building z-facility located at site  $k$ . Both  $y_j$  and  $z_k$  are binary value, which means whether building a y-facility at site  $j$  and z-facility at site  $k$ . The objective is to find two subsets  $\mathcal{C}^1 \subseteq \mathcal{J}$  and  $\mathcal{C}^2 \subseteq \mathcal{K}$  with minimal cost

This work was supported in part by the National Natural Science Foundation of China under grant U1834211, 61790575, U1934220, and in part by the Foundation of China Academy of Railway Sciences Corporation Limited under Grant 2019YJ071. (Corresponding author: Qi Zhang)

Shuxin Ding, Qi Zhang, and Zhiming Yuan are with the Signal and Communication Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China. dingshuxin@rails.cn; zhangqi@rails.cn; zhimingyuan@hotmail.com

$c(\mathcal{C}^1, \mathcal{C}^2) := \sum_{j \in \mathcal{C}^1} c_j^1 + \sum_{k \in \mathcal{C}^2} c_k^2$  covering all the demand nodes, i.e., for each demand node  $i \in \mathcal{I}$  there exists at least one y-facility  $j \in \mathcal{C}^1$  and z-facility  $k \in \mathcal{C}^2$  which ensures  $a_{ij} = 1$  and  $b_{ik} = 1$  simultaneously. A standard binary nonlinear programming formulation of TLCSCP is defined as

$$\min \sum_{j \in \mathcal{J}} c_j^1 y_j + \sum_{k \in \mathcal{K}} c_k^2 z_k \quad (4)$$

$$\text{s.t.} \quad \left( \sum_{j \in \mathcal{J}} a_{ij} y_j \right) \cdot \left( \sum_{k \in \mathcal{K}} b_{ik} z_k \right) \geq 1 \quad \forall i \in \mathcal{I} \quad (5)$$

$$y_j \in \{0, 1\} \quad \forall j \in \mathcal{J} \quad (6)$$

$$z_k \in \{0, 1\} \quad \forall k \in \mathcal{K}. \quad (7)$$

where (4) minimize the building cost of two kinds of facilities. (5) ensures that for each demand node, it is covered at least one y-facility and z-facility simultaneously. Equations (6) and (7) ensures decision variables are binary value. Since  $a_{ij}$ ,  $b_{ik}$ ,  $y_j$  and  $z_k$  are binary value, TLCSCP is equivalent to the following integer linear programming formulation:

$$\min \sum_{j \in \mathcal{J}} c_j^1 y_j + \sum_{k \in \mathcal{K}} c_k^2 z_k$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{J}} a_{ij} y_j \geq 1 \quad \forall i \in \mathcal{I} \quad (8)$$

$$\sum_{k \in \mathcal{K}} b_{ik} z_k \geq 1 \quad \forall i \in \mathcal{I} \quad (9)$$

$$y_j \in \{0, 1\} \quad \forall j \in \mathcal{J}$$

$$z_k \in \{0, 1\} \quad \forall k \in \mathcal{K},$$

where (8) and (9) linearize (5). And similar to SCP [13], TLCSCP is also a NP-hard combinatorial optimization problem.

Then, the Generalized Uncertain Two-Level Cooperative Set Covering Problem (GUTLCSCP) is formulated based on TLCSCP, which introduces uncertainty into covering model.  $a_{ij}$  and  $b_{ik}$  are independent random binary variable: with a probability of  $1 - p_{ij}$  when  $a_{ij} = 1$  and  $p_{ij}$  when  $a_{ij} = 0$ ; with a probability of  $1 - q_{ik}$  when  $b_{ik} = 1$  and  $q_{ik}$  when  $b_{ik} = 0$ . Since the probabilities are assumed to be independent, the probability of two sets  $\mathcal{C}^1$  and  $\mathcal{C}^2$  cooperatively covering demand node  $i$  is as follows:

$$P\left(\sum_{j \in \mathcal{C}^1} a_{ij} \geq 1\right) = 1 - \prod_{j \in \mathcal{C}^1} p_{ij}, \quad P\left(\sum_{k \in \mathcal{C}^2} b_{ik} \geq 1\right) = 1 - \prod_{k \in \mathcal{C}^2} q_{ik}.$$

$$P\left[\left(\sum_{j \in \mathcal{C}^1} a_{ij} \geq 1\right) \cdot \left(\sum_{k \in \mathcal{C}^2} b_{ik} \geq 1\right)\right] = \left(1 - \prod_{j \in \mathcal{C}^1} p_{ij}\right) \cdot \left(1 - \prod_{k \in \mathcal{C}^2} q_{ik}\right).$$

Then, the GUTLCSCP can be formulated as a binary model given by

$$\min \sum_{j \in \mathcal{J}} c_j^1 y_j + \sum_{k \in \mathcal{K}} c_k^2 z_k$$

$$\text{s.t.} \quad P\left(\sum_{j \in \mathcal{J}} a_{ij} y_j \geq 1, \sum_{k \in \mathcal{K}} b_{ik} z_k \geq 1\right) \geq \alpha \quad \forall i \in \mathcal{I} \quad (10)$$

$$y_j \in \{0, 1\} \quad \forall j \in \mathcal{J}$$

$$z_k \in \{0, 1\} \quad \forall k \in \mathcal{K}.$$

When a solution  $y^* \in \{0, 1\}^{n_1}$  and  $z^* \in \{0, 1\}^{n_2}$  is feasible for the GUTLCSCP, (10) is equivalent to

$$\left(1 - \prod_{j \in \mathcal{C}^1(y^*)} p_{ij}\right) \cdot \left(1 - \prod_{k \in \mathcal{C}^2(z^*)} q_{ik}\right) \geq \alpha \quad (11)$$

for all  $i \in \mathcal{I}$  with  $\mathcal{C}^1(y^*) = \{j \in \mathcal{J} | y_j^* = 1\}$  and  $\mathcal{C}^2(z^*) = \{k \in \mathcal{K} | z_k^* = 1\}$ . The sets  $\mathcal{C}^1(y^*)$  and  $\mathcal{C}^2(z^*)$  satisfying (11) are referred to as two-level-cooperative  $\alpha$ -cover. Note that these covering probabilities are assumed as independent variables (see [10]). A more realistic model with dependent probabilities may be further analyzed (out of scope).

Therefore, GUTLCSCP can be reformulated as:

$$\begin{aligned} \min \quad & \sum_{j \in \mathcal{J}} c_j^1 y_j + \sum_{k \in \mathcal{K}} c_k^2 z_k \\ \text{s.t.} \quad & \left(1 - \prod_{j \in \mathcal{J}} p_{ij}^{y_j}\right) \cdot \left(1 - \prod_{k \in \mathcal{K}} q_{ik}^{z_k}\right) \geq \alpha \quad \forall i \in \mathcal{I} \quad (12) \\ & y_j \in \{0, 1\} \quad \forall j \in \mathcal{J} \\ & z_k \in \{0, 1\} \quad \forall k \in \mathcal{K}. \end{aligned}$$

where (12) is a nonlinear constraint. A linear approximation method is given as follows.

In (12), set  $m_i = \prod_{j \in \mathcal{J}} p_{ij}^{y_j}$ ,  $n_i = \prod_{k \in \mathcal{K}} q_{ik}^{z_k}$ ,  $(1 - m_i)(1 - n_i) \geq \alpha$  for all  $i \in \mathcal{I}$ . The original constraint (12) can be reformulated as:

$$\begin{cases} m_i = \prod_{j \in \mathcal{J}} p_{ij}^{y_j} \\ n_i = \prod_{k \in \mathcal{K}} q_{ik}^{z_k} \\ (1 - m_i)(1 - n_i) \geq \alpha \end{cases} \iff \begin{cases} \ln(m_i) = \sum_{j \in \mathcal{J}} \ln(p_{ij}) y_j \\ \ln(n_i) = \sum_{k \in \mathcal{K}} \ln(q_{ik}) z_k \\ (1 - m_i)(1 - n_i) \geq \alpha \end{cases} \quad (13)$$

where for all  $i \in \mathcal{I}$  with  $m_i, n_i, \alpha \in [0, 1]$ . Therefore, GUTLCSCP can be reformulated as:

$$\begin{aligned} \min \quad & \sum_{j \in \mathcal{J}} c_j^1 y_j + \sum_{k \in \mathcal{K}} c_k^2 z_k \\ \text{s.t.} \quad & \ln(m_i) = \sum_{j \in \mathcal{J}} \ln(p_{ij}) y_j \quad \forall i \in \mathcal{I} \quad (14) \end{aligned}$$

$$\ln(n_i) = \sum_{k \in \mathcal{K}} \ln(q_{ik}) z_k \quad \forall i \in \mathcal{I} \quad (15)$$

$$(1 - m_i)(1 - n_i) \geq \alpha \quad \forall i \in \mathcal{I} \quad (16)$$

$$y_j \in \{0, 1\} \quad \forall j \in \mathcal{J}$$

$$z_k \in \{0, 1\} \quad \forall k \in \mathcal{K}$$

$$0 \leq m_i \leq 1 \quad \forall i \in \mathcal{I} \quad (17)$$

$$0 \leq n_i \leq 1 \quad \forall i \in \mathcal{I}. \quad (18)$$

The key issue is to deal with the constraint (16). Since  $m_i, n_i, \alpha \in [0, 1]$ , (16) can be reformulated in part as follows.

$$\begin{cases} 1 - m_i \geq \alpha \\ 1 - n_i \geq \alpha \end{cases} \iff \begin{cases} \sum_{j \in \mathcal{J}} \ln(p_{ij}) y_j \leq \ln(1 - \alpha) \\ \sum_{k \in \mathcal{K}} \ln(q_{ik}) z_k \leq \ln(1 - \alpha) \end{cases} \quad (19)$$

for all  $i \in \mathcal{I}$ .

According to the (14) and (15),  $\ln(m_i)$  and  $\ln(n_i)$  are linear functions with respect to  $y_j$  and  $z_k$ . Besides, we have obtained (19). As a result, we can construct constraints like

$$\beta \ln(m_i) + \gamma \ln(n_i) \leq \ln(F_i(\alpha, \beta, \gamma)), \quad (20)$$

where  $\beta + \gamma = 1$ ,  $F_i(\alpha, \beta, \gamma)$  is a function with respect to  $\alpha, \beta$  and  $\gamma$  for all  $i \in \mathcal{I}$ . In order to determine the parameters

of the constraint (20), we can find the function tangent to (16).

The constraint (20) can be reformulated as follows

$$\beta \ln(m_i) + \gamma \ln(n_i) \leq \ln(F_i(\alpha, \beta, \gamma)) \iff m_i^\beta n_i^\gamma \leq F_i(\alpha, \beta, \gamma). \quad (21)$$

Set  $(1 - m_i)(1 - n_i) = \alpha$ , then  $m_i = 1 - \alpha/(1 - n_i)$ . We can substitute it into (21) and obtain

$$f(n_i) = \left(1 - \frac{\alpha}{1 - n_i}\right)^\beta n_i^\gamma. \quad (22)$$

Then by determining the first derivative of (22), which is  $f'(n_i) = 0$ , the tangency function is obtained.

The solution to  $f'(n_i) = 0$  is as follows

$$\begin{cases} n_{i1} = 0 \\ n_{i2} = 1 - \alpha \\ n_{i3} = \frac{2\gamma + \alpha\beta - \alpha\gamma - \sqrt{\alpha(4\beta\gamma + \alpha\beta^2 + \alpha\gamma^2 - 2\alpha\beta\gamma)}}{2\gamma} \\ n_{i4} = \frac{2\gamma + \alpha\beta - \alpha\gamma + \sqrt{\alpha(4\beta\gamma + \alpha\beta^2 + \alpha\gamma^2 - 2\alpha\beta\gamma)}}{2\gamma} \end{cases}.$$

The corresponding tangency function is obtained when  $n_{i3}$  is selected. Fig. 1(a) shows comparison between constraint (20) after linear approximation and nonlinear constraint (16) when  $\alpha = 0.9$ ,  $\beta = 0.5$  and  $\gamma = 0.5$ . The range for x-axis and y-axis are determined by (19) within  $[0, 0.1]$ . Fig. 1(a) shows that there exists regions between these two constraints with one pair of  $\beta/\gamma$ . Therefore, multiple combinations of  $\beta/\gamma$  are needed. Fig. 1(b) shows the comparison when  $\beta = [0.1 \ 0.3 \ 0.5 \ 0.7 \ 0.9]$ . Fig. 1(c) combines these  $\beta/\gamma$  together to get an intersection of those constraints. Obviously, the linear approximate constraints show great similarity to the original nonlinear constraint. The decision space under the linear approximate constraints is slightly bigger than under nonlinear constraint. Relaxing the problem in (12) leads to the following linear approximation formulation of the GUTLCSCP (GUTLCSCP-LA):

$$\begin{aligned} \min \quad & \sum_{j \in \mathcal{J}} c_j^1 y_j + \sum_{k \in \mathcal{K}} c_k^2 z_k \\ \text{s.t.} \quad & \sum_{j \in \mathcal{J}} \ln(p_{ij}) y_j \leq \ln(1 - \alpha) \quad \forall i \in \mathcal{I} \\ & \sum_{k \in \mathcal{K}} \ln(q_{ik}) z_k \leq \ln(1 - \alpha) \quad \forall i \in \mathcal{I} \\ & \beta \sum_{j \in \mathcal{J}} \ln(p_{ij}) y_j + \gamma \sum_{k \in \mathcal{K}} \ln(q_{ik}) z_k \\ & \leq \ln \left[ \left(1 - \frac{\alpha}{1 - \delta}\right)^\beta n_i^\gamma \right] \quad \forall i \in \mathcal{I} \\ & y_j \in \{0, 1\} \quad \forall j \in \mathcal{J} \\ & z_k \in \{0, 1\} \quad \forall k \in \mathcal{K}, \end{aligned}$$

where  $\delta = \frac{2\gamma + \alpha\beta - \alpha\gamma - \sqrt{\alpha(4\beta\gamma + \alpha\beta^2 + \alpha\gamma^2 - 2\alpha\beta\gamma)}}{2\gamma}$ .  $\beta, \gamma \in [0, 1]$  are constants or vectors with  $\beta + \gamma = 1$ .

*Remark 1:* The linear approximation (LA) method used here actually transforms the original problem to a constraint relaxed problem as an integer linear programming program. All the problems in this paper with LA are the constraint relaxed version of the original problems.

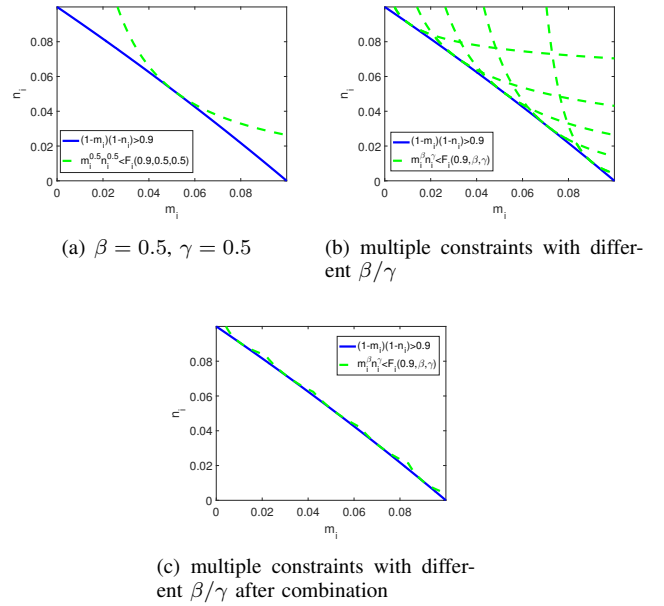


Fig. 1. Comparison between constraint (20) after linear approximation and nonlinear constraint (16) when  $\alpha = 0.9$ .

## B. Modeling the Robust Uncertain Two-Level Cooperative Set Covering Problem

The Robust Uncertain Two-Level Cooperative Set Covering Problem (RUTLCSCP) is formulated based on the GUTLCSCP, with fluctuation of the probabilities  $p_{ij}$  and  $q_{ik}$ .

In real-world applications, the probabilities  $p_{ij}$  and  $q_{ik}$  are not precisely known [13]. They can be estimated based on historical data. However, these estimated values could not reflect the whole situation. In some situations, estimated values may be too optimistic, while in other situations, they may be too pessimistic. Hence, there exists a natural fluctuation of the probabilities. Therefore, in order to model the effect of these fluctuations, an interval is established based on the *nominal value*. This interval covers the range of probabilities. As a result, this description of probability is more reasonable than using a particular value [13].

The following description is based on a  $\Gamma$ -scenario set proposed by Bertsimas and Sim [14]. There are at most  $\Gamma$  values deviate from their *nominal value*. When  $\Gamma = n$ , all parameters are allowed to deviate, which is equivalent to Soyster's robust model [12]. However, this model is too conservative.  $\Gamma$  models the risk attitude of the parameters [13], and it is also called the budget of uncertainty.

We assume that  $p_{ij}$  and  $q_{ik}$  are uncertain variable within the interval  $[\bar{p}_{ij}, \bar{p}_{ij} + \hat{p}_{ij}] \subseteq [0, 1]$  and  $[\bar{q}_{ik}, \bar{q}_{ik} + \hat{q}_{ik}] \subseteq [0, 1]$  where  $\bar{p}_{ij} \geq 0$  and  $\bar{q}_{ik} \geq 0$  are the *nominal value*,  $\hat{p}_{ij} \geq 0$  and  $\hat{q}_{ik} \geq 0$  are the *worst case deviation*. The two  $\Gamma$ -scenario sets are given by

$$\begin{aligned} \mathcal{U}_1^{\Gamma_i} &:= \left\{ p_i : \forall j \in \mathcal{J} : p_{ij} \in [\bar{p}_{ij}, \bar{p}_{ij} + \hat{p}_{ij}], \sum_{j \in \mathcal{J}} \frac{p_{ij} - \bar{p}_{ij}}{\hat{p}_{ij}} \leq \Gamma_i \right\} \\ \mathcal{U}_2^{\Gamma_i} &:= \left\{ q_i : \forall k \in \mathcal{K} : q_{ik} \in [\bar{q}_{ik}, \bar{q}_{ik} + \hat{q}_{ik}], \sum_{k \in \mathcal{K}} \frac{q_{ik} - \bar{q}_{ik}}{\hat{q}_{ik}} \leq \Gamma_i \right\}, \end{aligned}$$

for all  $i \in \mathcal{I}$ , where  $p_i := (p_{ij})_{j \in \mathcal{J}}$ ,  $q_i := (q_{ik})_{k \in \mathcal{K}}$ .

The difference between RUTLCSCP and GUTLCSCP is that for any  $i \in \mathcal{I}$ , there exists two-level-cooperative  $\alpha$ -cover in RUTLCSCP with probabilities satisfying  $p_{i\cdot} \in \mathcal{U}_1^{\Gamma_i}$  and  $q_{i\cdot} \in \mathcal{U}_2^{\Gamma_i}$ . We can consider the worst case: there exists  $\Gamma_i \in \mathbb{N}_0$  entries in  $p_{i\cdot}$  and  $q_{i\cdot}$  derive from their nominal value, which are the worst case deviation. The other entries in  $p_{i\cdot}$  and  $q_{i\cdot}$  are their nominal values  $\bar{p}_{ij}$  and  $\bar{q}_{ik}$ . A  $\Gamma$ -robust two-level-cooperative  $\alpha$ -cover is defined as follows.

**Definition 1:** ( $\Gamma$ -robust two-level-cooperative  $\alpha$ -cover). Set  $i \in \mathcal{I}$ ,  $\Gamma_i \in \mathbb{N}_0$ ,  $\Gamma = (\Gamma_i)_{i \in \mathcal{I}}$ ,  $\alpha \in [0, 1)$ . For all  $j \in \mathcal{J}$  and  $k \in \mathcal{K}$ ,  $p_{ij}$  are within range  $[\bar{p}_{ij}, \bar{p}_{ij} + \hat{p}_{ij}] \subseteq [0, 1]$ ,  $q_{ik}$  are within range  $[\bar{q}_{ik}, \bar{q}_{ik} + \hat{q}_{ik}] \subseteq [0, 1]$ . The worst-case coverage probability for set  $C^1 \subseteq \mathcal{J}$  and set  $C^2 \subseteq \mathcal{K}$  can be defined by

$$P_{\Gamma_i} \left( \sum_{j \in C^1} a_{ij} \geq 1 \right) := 1 - \max_{\{\mathcal{U}_1 \subseteq C^1: |\mathcal{U}_1| \leq \Gamma_i\}} \left\{ \prod_{j \in \mathcal{U}_1} (\bar{p}_{ij} + \hat{p}_{ij}) \cdot \prod_{j \in C^1 \setminus \mathcal{U}_1} \bar{p}_{ij} \right\}$$

$$P_{\Gamma_i} \left( \sum_{k \in C^2} b_{ik} \geq 1 \right) := 1 - \max_{\{\mathcal{U}_2 \subseteq C^2: |\mathcal{U}_2| \leq \Gamma_i\}} \left\{ \prod_{k \in \mathcal{U}_2} (\bar{q}_{ik} + \hat{q}_{ik}) \cdot \prod_{k \in C^2 \setminus \mathcal{U}_2} \bar{q}_{ik} \right\}.$$

A  $\Gamma$ -robust two-level-cooperative  $\alpha$ -cover with  $C^1 \subseteq \mathcal{J}$  and  $C^2 \subseteq \mathcal{K}$  have a worst-case coverage probability  $P_{\Gamma_i} \left( \sum_{j \in C^1} a_{ij} \geq 1 \right) \cdot P_{\Gamma_i} \left( \sum_{k \in C^2} b_{ik} \geq 1 \right)$  greater or equals to  $\alpha$ . When all  $i \in \mathcal{I}$  for set  $C^1$  and set  $C^2$  satisfying  $\Gamma$ -robust two-level-cooperative  $\alpha$ -cover, then a  $\Gamma$ -robust two-level-cooperative  $\alpha$ -cover is obtained.

The RUTLCSCP is to find a  $\Gamma$ -robust two-level-cooperative  $\alpha$ -cover of minimum costs. A nonlinear formulation can be defined in the following:

$$\begin{aligned} \min \quad & \sum_{j \in \mathcal{J}} c_j^1 y_j + \sum_{k \in \mathcal{K}} c_k^2 z_k \\ \text{s.t.} \quad & P_{\Gamma_i} \left( \sum_{j \in \mathcal{J}} a_{ij} y_j \geq 1 \right) \cdot P_{\Gamma_i} \left( \sum_{k \in \mathcal{K}} b_{ik} z_k \geq 1 \right) \geq \alpha \quad \forall i \in \mathcal{I} \\ & y_j \in \{0, 1\} \quad \forall j \in \mathcal{J} \\ & z_k \in \{0, 1\} \quad \forall k \in \mathcal{K}. \end{aligned} \quad (23)$$

A solution  $y^* \in \{0, 1\}^{n_1}$ ,  $z^* \in \{0, 1\}^{n_2}$  is called robust feasible when  $\Gamma$ -robust two-level-cooperative  $\alpha$ -cover is satisfied. There exists two maximum subproblems in (23) defined as

$$\beta_i^1(y, \Gamma_i) := \max_{\{\mathcal{U}_1 \subseteq C^1(y): |\mathcal{U}_1| \leq \Gamma_i\}} \left\{ \prod_{j \in \mathcal{U}_1} (\bar{p}_{ij} + \hat{p}_{ij})^{y_j} \cdot \prod_{j \in \mathcal{J} \setminus \mathcal{U}_1} \bar{p}_{ij}^{y_j} \right\} \quad (24)$$

$$\beta_i^2(z, \Gamma_i) := \max_{\{\mathcal{U}_2 \subseteq C^2(z): |\mathcal{U}_2| \leq \Gamma_i\}} \left\{ \prod_{k \in \mathcal{U}_2} (\bar{q}_{ik} + \hat{q}_{ik})^{z_k} \cdot \prod_{k \in \mathcal{K} \setminus \mathcal{U}_2} \bar{q}_{ik}^{z_k} \right\} \quad (25)$$

where for all  $i \in \mathcal{I}$ .

Therefore, the RUTLCSCP can be reformulated as

$$\begin{aligned} \min \quad & \sum_{j \in \mathcal{J}} c_j^1 y_j + \sum_{k \in \mathcal{K}} c_k^2 z_k \\ \text{s.t.} \quad & [1 - \beta_i^1(y, \Gamma_i)] \cdot [1 - \beta_i^2(z, \Gamma_i)] \geq \alpha \quad \forall i \in \mathcal{I} \\ & y_j \in \{0, 1\} \quad \forall j \in \mathcal{J} \\ & z_k \in \{0, 1\} \quad \forall k \in \mathcal{K}. \end{aligned} \quad (26)$$

Similarly, we can develop the linear approximate model of the RUTLCSCP based on the GUTLCSCP-LA. Meanwhile, applying the strong duality theorem, we can develop the robust counterpart (RC) of the robust model RUTLCSCP-LA-RC, which is a compact mixed-integer linear programming problem:

$$\begin{aligned} \min \quad & \sum_{j \in \mathcal{J}} c_j^1 y_j + \sum_{k \in \mathcal{K}} c_k^2 z_k \\ \text{s.t.} \quad & \sum_{j \in \mathcal{J}} \ln(\bar{p}_{ij}) y_j + \sum_{j \in \mathcal{J}} \zeta_{ij}^1 + \Gamma_i \eta_i^1 \leq \ln(1 - \alpha) \quad \forall i \in \mathcal{I} \\ & \sum_{k \in \mathcal{K}} \ln(\bar{q}_{ik}) z_k + \sum_{k \in \mathcal{K}} \zeta_{ik}^2 + \Gamma_i \eta_i^2 \leq \ln(1 - \alpha) \quad \forall i \in \mathcal{I} \\ & \beta \left[ \sum_{j \in \mathcal{J}} \ln(\bar{p}_{ij}) y_j + \sum_{j \in \mathcal{J}} \zeta_{ij}^1 + \Gamma_i \eta_i^1 \right] \\ & + \gamma \left[ \sum_{k \in \mathcal{K}} \ln(\bar{q}_{ik}) z_k + \sum_{k \in \mathcal{K}} \zeta_{ik}^2 + \Gamma_i \eta_i^2 \right] \\ & \leq \ln \left[ \left( 1 - \frac{\alpha}{1 - \delta} \right)^\beta n_i^\gamma \right] \quad \forall i \in \mathcal{I} \\ & \zeta_{ij}^1 + \eta_i^1 \geq (\ln(\bar{p}_{ij} + \hat{p}_{ij}) - \ln(\bar{p}_{ij})) y_j \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \\ & \zeta_{ik}^2 + \eta_i^2 \geq (\ln(\bar{q}_{ik} + \hat{q}_{ik}) - \ln(\bar{q}_{ik})) z_k \quad \forall i \in \mathcal{I}, k \in \mathcal{K} \\ & \zeta_{ij}^1 \geq 0 \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \\ & \zeta_{ik}^2 \geq 0 \quad \forall i \in \mathcal{I}, k \in \mathcal{K} \\ & \eta_i^1 \geq 0 \quad \forall i \in \mathcal{I} \\ & \eta_i^2 \geq 0 \quad \forall i \in \mathcal{I} \\ & y_j \in \{0, 1\} \quad \forall j \in \mathcal{J} \\ & z_k \in \{0, 1\} \quad \forall k \in \mathcal{K}, \end{aligned}$$

where  $\delta = \frac{2\gamma + \alpha\beta - \alpha\gamma - \sqrt{\alpha(4\beta\gamma + \alpha\beta^2 + \alpha\gamma^2 - 2\alpha\beta\gamma)}}{2\gamma}$ .  $\beta, \gamma \in [0, 1]$  are constants or vectors with  $\beta + \gamma = 1$ .

### III. PROPERTIES OF THE MODEL

There exist nonlinear, noncompact constraints, and maximum subproblems in the proposed RUTLCSCP, which are hard to solve. A definition and two propositions are provided as follows.

**Definition 2:** ( $\varepsilon$ -under-approximate solution). Given a scalar  $\varepsilon > 0$ , a  $\varepsilon$ -under-approximate solution has a larger feasible region with constraints relaxed than the original feasible region with the original constraints. The new feasible region is obtained by linear approximation of the nonlinear constraints, i.e.,  $X_{LA} \in \Omega_{LA} = \{x | C_i(x)(1 + \varepsilon) \geq \alpha, x = \arg \min_{X \in \Omega} \{F(X)\}, i \in \mathcal{I}\}$ , where  $\Omega$  is the feasible region of the original problem and  $\Omega_{LA}$  is the approximate feasible region.

**Proposition 1:** Suppose the solution to the linear approximate problem is  $X_{LA}$  with the objective value  $F_{LA}(X_{LA})$ , while the solution for the nonlinear constraints problem is  $X$  with the objective value  $F(X)$ . Then we will have  $F_{LA}(X_{LA}) \leq F(X)$ , which is a lower bound on the optimal objective function. If the nonlinear constraints are satisfied when we substitute the solution  $X_{LA}$  into the original problem with nonlinear constraints, we will have  $F_{LA}(X_{LA}) = F(X)$ . The nonlinear constraints problems include the GUTLCSCP and the RUTLCSCP, while the

linear approximate problems are the GUTLCSCP-LA and the RUTLCSCP-LA-RC.

*Proof:* The solution to the linear approximate problem is  $X_{LA} \in \Omega_{LA}$ , while the solution for the nonlinear constraints problem is  $X \in \Omega$ . According to Fig. 1(c), nonlinear constraints are relaxed by the linear approximation method. Therefore,  $\Omega \in \Omega_{LA}$  is a subset of the approximate feasible region. As a result,  $F_{LA}(X_{LA}) \leq F(X)$ . When the solution  $X_{LA}$  satisfies the nonlinear constraints, that means  $X_{LA} \in \Omega$ . Therefore, we will have  $F_{LA}(X_{LA}) = F(X)$ . ■

*Proposition 2:* If the problem after linear approximation (GUTLCSCP-LA, RUTLCSCP-LA-RC) has no solution, the original problem with nonlinear constraints (RUTLCSCP-LA-RC, RUTLCSCP) has no solution as well.

*Proof:* Based on Proposition 1, we have  $\Omega \subseteq \Omega_{LA}$ . If there is no solution in the feasible region  $\Omega_{LA}$ , then there is no solution in the feasible region  $\Omega$  as well. In other words, if there is no solution in the linear approximate problem, there is no solution in the original problem with nonlinear constraints. ■

Therefore, based on the above propositions, as for problems in different scales, we could use an exact method or solver (e.g., IBM-ILOG-CPLEX) to solve the RUTLCSCP-LA-RC in order to obtain the exact solution to the RUTLCSCP if the equality condition in Proposition 1 is met. Otherwise,  $\varepsilon$ -under-approximate solution is obtained.

#### IV. COMPUTATIONAL EXPERIMENTS AND ANALYSIS

This section is devoted to the performance investigation of the proposed model. At first, we present a RUTLCSCP test-case generator, which can produce instances of different scales. Then, we solve the problem, which includes exact solutions for RUTLCSCP-LA-RC and approximate solutions for RUTLCSCP. All experiments were carried out on a PC with Intel Xeon E5 CPU 2.60GHz and 64 GB internal memory. RUTLCSCP-LA-RC problems were implemented in MATLAB R2016a using YALMIP as the modeling language and CPLEX 12.5 with default parameter settings.

##### A. Test-Case Generator

Due to the lack of benchmark instances for the RUTLCSCP in literature, we consider the following parameter setting. The fixed costs coefficients building y-facility  $c_j^1$  and z-facility  $c_k^2$  were both randomly generated by sampling from a uniform distribution in  $[0, 100]$ . The nominal value of probabilities  $\bar{p}_{ij}$  and  $\bar{q}_{ik}$  were both obtained by sampling from a uniform distribution in  $[0.9, 1.0]$ . Deviations for the default probability  $\hat{p}_{ij}$  and  $\hat{q}_{ik}$  were both taken from a uniform distribution in  $[0, 0.1]$ . Besides, we consider two covering ranges,  $yr$  and  $zr$ , for these two kinds of facilities. If the Euclidean distance of the demand node and facility location is greater than the covering range, the corresponding probability  $p_{ij}$  or  $q_{ik}$  is 0. Each demand node serves as a candidate location site for y-facility and z-facility, i.e.,  $\mathcal{I} = \mathcal{J} = \mathcal{K}$ . The positions of the demand nodes are randomly generated within the region  $Ax \times Ay$ . All the RUTLCSCP formulations were solved for the parameters

$\alpha \in \{0.8, 0.85, 0.9\}$  and  $\Gamma \in \{0, \dots, |\mathcal{I}|\}$ . 10 cases were considered. For each case, we randomly generated five different instances. In total, 10125 derived RUTLCSCP instances were generated. The detailed information of these instances is in Table I.

TABLE I  
THE TEST-CASE FOR RUTLCSCP

Instance	$( \mathcal{I} ,  \mathcal{J} ,  \mathcal{K} )$	$(yr/km, zr/km)$	$(Ax/km, Ay/km)$
P1.1–P1.5	(20, 20, 20)	(10, 5)	(25, 25)
P2.1–P2.5	(25, 25, 25)	(10, 5)	(25, 25)
P3.1–P3.5	(30, 30, 30)	(10, 5)	(25, 25)
P4.1–P4.5	(40, 40, 40)	(14, 7)	(50, 50)
P5.1–P5.5	(50, 50, 50)	(14, 7)	(50, 50)
P6.1–P6.5	(60, 60, 60)	(14, 7)	(50, 50)
P7.1–P7.5	(80, 80, 80)	(20, 10)	(100, 100)
P8.1–P8.5	(100, 100, 100)	(20, 10)	(100, 100)
P9.1–P9.5	(120, 120, 120)	(20, 10)	(100, 100)
P10.1–P10.5	(140, 140, 140)	(20, 10)	(100, 100)

##### B. Results and Analysis

We found that the approximation accuracy of the constraints is related to the amount of the  $\beta/\gamma$  pairs. If we use more pairs, the approximation will be better, which increases the total running time of the algorithm. Therefore, one needs to balance these two conflicts. Here we considered the combination  $\beta = [0.001 \ 0.01 \ 0.05 \ 0.1 \ 0.15 \ 0.2 \ 0.3 \ 0.4 \ 0.5 \ 0.6 \ 0.7 \ 0.8 \ 0.85 \ 0.9 \ 0.95 \ 0.99 \ 0.999]$  based on empirical testing, where  $\gamma = 1 - \beta$ .

The results for RUTLCSCP are presented in Table II with the following statistics:

- *Proof of opt:* The proportion of instances in which the solution was proven to be optimal.
- *Time:* Arithmetic mean of run times in seconds.
- *CV (constraint violation):* The proportion of violated constraints in RUTLCSCP with feasible  $\varepsilon$ -under-approximate solution.
- *Degree of feasibility:* The ratio of feasible solutions without any violated constraint in RUTLCSCP-LA-RC and the total number of instances.

From Table II, there exists infeasible solutions for RUTLCSCP-LA-RC since the degree of feasibility is less than 100%. These instances are especially those with  $\alpha = 0.9$  and  $\Gamma \geq 1$ . As a result, the corresponding instances of RUTLCSCP have no solution. Besides, for some instances, the solutions violate the original nonlinear constraints but feasible to RUTLCSCP-LA-RC, which are the  $\varepsilon$ -under-approximate solutions. These instances are shown in Table III with 12 instances types.  $\phi$  represents the total constraint violations and  $\#$  stands for the proportion of violations with total nonlinear constraints. The solutions for the remaining instances are also solutions for the original problem, RUTLCSCP. Most of the  $\phi$  of the approximate solutions are at a level of E-4~E-6, and with only one violated constraint, which means great approximation. The corresponding objective value is closely lower than the optimal value, which is an efficient under-approximation and a lower bound.

For  $\alpha = 0.8$  and  $\alpha = 0.85$ , the objective values are the same as  $\Gamma \geq 1$ . However, for  $\alpha = 0.9$ , the objective values are different under different  $\Gamma$ . Most instances have the same objective value when  $\Gamma \geq 2$ . Note that in P8.3 ( $\alpha = 0.9$ ,

TABLE II  
COMPUTATIONAL RESULTS FOR RUTLCSCP

Instance	$\alpha = 0.8^*$			$\alpha = 0.85^*$			$\alpha = 0.9$			
	Opt. (%)	Time	CV (%)	Opt. (%)	Time	CV (%)	Opt. (%)	Time	CV (%)	Degree of feasibility (%)
P1.1–P1.5	100.00	0.14	0.00	100.00	0.17	0.00	100.00	0.10	0.00	61.90
P2.1–P2.5	100.00	0.21	0.00	100.00	0.25	0.00	98.75	0.20	0.05	61.54
P3.1–P3.5	100.00	0.24	0.00	100.00	0.33	0.00	99.20	0.37	0.03	80.65
P4.1–P4.5	100.00	0.32	0.00	100.00	0.48	0.00	100.00	0.30	0.00	21.95
P5.1–P5.5	100.00	0.56	0.00	100.00	0.76	0.00	100.00	0.44	0.00	41.18
P6.1–P6.5	100.00	1.03	0.00	100.00	1.01	0.00	100.00	0.44	0.00	21.31
P7.1–P7.5	80.25	1.48	0.25	100.00	1.60	0.00	100.00	0.74	0.00	1.23
P8.1–P8.5	100.00	3.00	0.00	80.00	3.76	0.20	96.10	2.81	0.04	40.59
P9.1–P9.5	100.00	4.59	0.00	100.00	5.98	0.00	99.18	4.58	0.01	40.50
P10.1–P10.5	100.00	7.29	0.00	80.14	8.33	0.14	100.00	5.68	0.00	20.57

\* Degree of feasibility are 100%.

TABLE III  
INSTANCE TYPES WITH CONSTRAINT VIOLATION

Instance	$\alpha$	$\Gamma$	Obj.	$\phi$	#
P2.2	0.9	0	463.94	7.40E-06	1/20
P3.1	0.9	0	309.00	1.29E-05	1/25
P7.1	0.8	1+	1464.12	3.79E-04	1/80
P8.1	0.9	0	1258.93	1.13E-05	1/100
P8.2	0.85	0	1514.56	4.21E-04	1/100
P8.2	0.9	0	1603.00	8.01E-06	1/100
P8.3	0.85	1+	1444.47	2.51E-04 <sup>‡</sup>	1/100
P8.3	0.9	0	1409.12	1.93E-04	1/100
P8.3	0.9	2+	1942.41 <sup>†</sup>	2.12E-04 <sup>‡</sup>	1/100
P9.2	0.9	0	1501.84	3.91E-04	1/120
P9.3	0.9	0	1312.91	3.97E-06	1/120
P10.3	0.85	1+	1408.94	3.44E-04	1/140

<sup>†</sup> The objective values are still varying as different  $\Gamma$ .

<sup>‡</sup> The total constraint violations are varying with different  $\Gamma$ .

$\Gamma \geq 2$ ) marked with <sup>†</sup>, the objective values are still varying as  $\Gamma \geq 2$ . When  $\Gamma = 2, 3, 6, 18, 19$ , the corresponding objective values are 1942.41. The values for the rest are 1947.82. The total constraint violations marked with <sup>‡</sup> means that they vary with different  $\Gamma$ . For example, in P8.3 ( $\alpha = 0.85$ ,  $\Gamma \geq 1$ ),  $\phi = 2.51E - 04$  when  $\Gamma = 1$ ; while  $\phi = 6.30E - 04$  when  $\Gamma \geq 2$ . In P8.3 ( $\alpha = 0.9$ ,  $\Gamma \geq 2$ ),  $\phi = 2.12E - 04$  when  $\Gamma = 2, 3, 6, 18, 19$ ; while the rest are constraint satisfied. Note that CPLEX can efficiently solve RUTLCSCP-LA-RC, with the computation time less than 10 seconds.

In summary, a set of 10125 instances are generated and solved with good quality and acceptable time. Up to 74.10% (7502 instances) are solved to optimality, 3.29% (333 instances) are under-approximation, and 22.62% (2290 instances) are with no solution.

## V. CONCLUSION

In this paper, we consider an extension of the set covering problem (SCP) called the robust uncertain two-level cooperative set covering problem (RUTLCSCP) by the integration of uncertainty in covering demand nodes. The concepts of probabilistic, robust optimization, and cooperative covering are combined. A compact mixed-integer linear programming (MILP) formulation for the RUTLCSCP is proposed. Computational experiments demonstrate that the RUTLCSCP can be efficiently solved with optimal solutions and a few under-approximate solutions.

In the future, over-approximation with more constraints and less feasible regions are likely to be investigated. Be-

sides, new exact or heuristic algorithms, new reformulation, and multi-level of the model can be considered. Meanwhile, the proposed model can be applied in many real-world applications, e.g., collaborative task assignment [11], joint allocation of heterogeneous stochastic resources [15], etc.

## REFERENCES

- [1] V. Chvatal, "A greedy heuristic for the set-covering problem," *Mathematics of Operations Research*, vol. 4, no. 3, pp. 233–235, 1979.
- [2] R. Z. Farahani, N. Asgari, N. Heidari, M. Hosseini, and M. Goh. "Covering problems in facility location: A review," *Computers & Industrial Engineering*, vol. 62, no. 1, pp. 368–407, 2012.
- [3] M. S. Daskin, "A maximum expected covering location model: formulation, properties and heuristic solution," *Transportation Science*, vol. 17, no. 1, pp. 48–70, 1983.
- [4] S. Ding, C. Chen, J. Chen, and B. Xin, "An improved particle swarm optimization deployment for wireless sensor networks," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 18, no. 2, pp. 107–112, 2014.
- [5] S. Ding, C. Chen, B. Xin, and J. Chen, "Status and progress in deployment optimization of firepower units," *Kongzhi Lilun Yu Yingyong/Control Theory and Applications*, vol. 32, no. 12, pp. 1569–1581, 2015.
- [6] P. Beraldi, and A. Ruszczyński, "The probabilistic set-covering problem," *Operations Research*, vol. 50, no. 6, pp. 956–967, 2002.
- [7] K. Aardal, M. Labbé, J. Leung, and M. Queyranne, "On the two-level uncapacitated facility location problem," *INFORMS Journal on Computing*, vol. 8, no. 3, pp. 289–301, 1996.
- [8] O. Berman, Z. Drezner, and D. Krass, "Cooperative cover location problems: the planar case," *IIE Transactions*, vol. 42, no. 3, pp. 232–246, 2009.
- [9] J. Pereira, and I. Averbakh, "The robust set covering problem with interval data," *Annals of Operations Research*, vol. 207, no. 1, pp. 217–235, 2013.
- [10] B. Xin, Y. Wang, and J. Chen, "An efficient marginal-return-based constructive heuristic to solve the sensor-weapon-target assignment problem," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 12, pp. 2536–2547, 2018.
- [11] W. Xu, C. Chen, S. Ding, and P. M. Pardalos, "A bi-objective dynamic collaborative task assignment under uncertainty using modified MOEA/D with heuristic initialization," *Expert Systems with Applications*, vol. 140, p. 112844, 2020.
- [12] A. L. Soyster, "Convex programming with set-inclusive constraints and applications to inexact linear programming," *Operations Research*, vol. 21, no. 5, pp. 1154–1157, 1973.
- [13] P. Lutter, D. Degel, C. Büsing, A. M. Koster, and B. Werners, "Improved handling of uncertainty and robustness in set covering problems," *European Journal of Operational Research*, vol. 263, no. 1, pp. 35–49, 2017.
- [14] D. Bertsimas and M. Sim, "The price of robustness," *Operations Research*, vol. 52, no. 1, pp. 35–53, 2004.
- [15] Y. Wang, B. Xin, L. Dou, and Z. Peng. "A heuristic initialized memetic algorithm for the joint allocation of heterogeneous stochastic resources," in *2019 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2019, pp. 1929–1936.